

AD-A133 766

UNTUTORED CONCEPT LEARNING(U) ILLINOIS UNIV AT URBANA
COORDINATED SCIENCE LAB G DEJONG JUN 83
AFOSR-TR-83-0771 F49620-82-K-0009

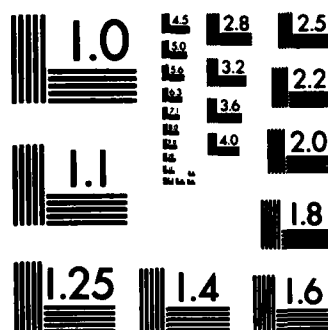
1/1

UNCLASSIFIED

F/G 6/4

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR-88-0771	2. GOVT ACCESSION NO. AD-A133766	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) "UNTUTORED CONCEPT LEARNING"		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL
7. AUTHOR(s) Gerald DeJong		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Illinois Urbana, IL 61801		8. CONTRACT OR GRANT NUMBER(s) F49620-82-K-0009
11. CONTROLLING OFFICE NAME AND ADDRESS AFOSR/NM Bolling AFB DC 20332		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE61102F; 2304/A2
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE June 1983
		13. NUMBER OF PAGES 7
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) <p>This paper discusses <u>explanatory schema acquisition</u>, a learning technique with several interesting properties. It does not require a teacher or concept matching predicate to be provided. It does not rely on searching a concept space to produce generalizations. It can acquire a new concept based on only one input example, although later inputs might result in refinement of learned concepts. These features are made possible by taking a very knowledge-based approach. The concepts that are learned are problem-solving.</p>		

DTIC FILE COPY

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

↑

Accession No. _____
 Date Recd _____ X
 By _____
 Title _____
 Author _____
 Subject _____
 Notes _____
 Price _____
 Source _____
 Remarks _____

A



SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Untutored Concept Learning
 Gerald DeJong
 Coordinated Science Laboratory
 1101 West Springfield Avenue
 University of Illinois
 Urbana, IL 61801
 (217) 333-0491

ABSTRACT

This paper discusses explanatory schema acquisition, a learning technique with several interesting properties. It does not require a teacher or concept matching predicate to be provided. It does not rely on searching a concept space to produce generalizations. It can acquire a new concept based on only one input example, although later inputs might result in refinement of learned concepts. These features are made possible by taking a very knowledge-based approach. The concepts that are learned are problem-solving schemata. Thus, the technique is not applicable to all types of learning. However, it provides a unique perspective on a large and interesting class of learning.

Subfield: Learning and Knowledge Acquisition
 Category: Short Paper
 Text Length: 2217 words

1. Introduction

This paper gives an overview of a learning technique being developed at the University of Illinois. The technique, called explanatory schema acquisition has some interesting properties. For example, it does not require a teacher or other oracle to select important examples; it is capable of one trial learning; and contrary to Mitchell's recent taxonomy of learning systems [14], it does little or no searching in the process of acquiring a new concept.

Before describing the technique we will pause briefly to consider what we might call the "standard theory of concept formation." This approach underlies much of the concept learning work in psychology and AI. In the standard theory, a system is given a number of inputs. Each input is composed of a set of features. A concept is a conjunctive and/or disjunctive combination of features. An input with the proper combination of features is an instance of the concept; otherwise it is not an instance. A teacher, usually a human, supplies sample inputs to the system together with the information of which concept (if any) the input is a concept of. The system's task is to discover the combination of features that compose each concept. This approach has been fruitfully applied to many diverse domains (for example, [12], [13], [21]) and is a cornerstone of the field of inductive inference.

This work was supported in part by the Air Force Office of Scientific Research under grant F49620-82-K-0009 and in part by the National Science Foundation under grant NSF IST 81-20254.

Approved for public release;
 distribution unlimited.

83 10 04 248

Now we can ask how we might construct an untutored concept learning system. At first this seems a bit of a non sequitur. Removing the teacher appears to result in no learning at all. Also the notion forming a concept from a single input seems suspicious.

The key is to adopt a much more knowledge-based approach. The learning algorithm to be described requires access to a large amount of domain knowledge. It is through reconciling a new input to the domain knowledge that learning and generalization occurs.

This is NOT to say that the proposed learning technique is domain specific. Techniques specific to a particular domain would be ad hoc and of very limited theoretical interest. Rather, explanatory schema acquisition is domain independent. Indeed, it has already been applied to three very different domains. The approach does, however, require access to a rich domain model. It is interaction with this rich domain information that determines whether or not concept acquisition is possible or desirable for a new input. The interaction also guides the generalization process.

2. Explanatory Schema Acquisition

The process involves three logically distinct (but possibly temporally concurrent) processes:

- 1) The new input is understood.
- 2) The input is evaluated to see if schema formation is warranted.
- 3) The input is generalized to a new schema.

For illustrative purposes we will assume that the input is the following brief natural language story. The assumption of natural language input is not necessary and, indeed, one of the current applications involves robot arm planning which, of course, is non-linguistic.

John, a bank teller, discovered that his boss, Fred, had embezzled \$100,000. John sent Fred an inter-office memo saying that he would inform the police unless he was given \$15,000. Fred paid John the money.

2.1. Understanding the Input

The requirements on the understanding process are not controversial. By "understanding" we mean nothing more than constructing a causally complete representation of the input. This requires that any crucial information missing from the story must be inferred and that the causal relations between components be discovered and made explicit. While this is not an easy task, it is one which has been the focus of a good deal of research, particularly for natural language texts ([1], [3], [10], [17], [19]).

We require that our representation have one component that is not generally included by understanding systems. We require that the understander maintain data dependency links ([4], [6]) justifying each element in the representation. The links connect each representation event with all of the inference rules from the domain model that were used to justify the event during the understanding process. This includes all causal information, goal enablements, planning information, etc. This makes explicit in the final story representation the reasons the system had for connecting events in a particular way. For example, in the blackmail story John's demanding that Fred give him \$15,000 is explicitly mentioned. The system must infer

that John has the goal of possessing the \$15,000. This is a necessary inference. A system cannot be said to have "understood" the input (in any sense of the word) if it does not make this inference.

We require not only that the inference be made but that it be justified by including data dependency links to the appropriate inference rules. In this example, the relevant inference rules state that all volitional actions are done in service of goals and that any kind of request is probably done in service of the goal of possessing the requested object.

By and large, current understanding systems do not include these backpointers to inference rules in the final representation. References to the inference rules exist, rather, only in a trace of the understander's processing. In most current systems they are available but simply not included. We will insist that they also be explicitly stated in the understood representation. We call the amalgam of all of these data dependency links the Inference Justification Network.

2.2. To Generalize or Not to Generalize

There are five aspects to be considered when deciding whether or not to generalize an input into a new schema. By hypothesis we will assume that the input did not match an existing schema (if it had then the system already possesses the desired schema and indeed that schema would have been used to process the story). If any of these five conditions does not hold, constructing a new schema from this input is inappropriate.

The criteria are:

- 1) Is the main goal of a character achieved?
- 2) Is the goal a general one?
- 3) Are the resources required by the goal achiever generally available?
- 4) Is this new method of achieving the goal at least as effective as the other known volitional schemata to achieve this goal?
- 5) Does the input match one of the known generalizedable patterns?

These criteria are tested for all goals in the story. The first criterion "Was the goal achieved?" is self explanatory and easily judged. The second "Is it a general goal?" and the third "Are the resources generally available?" require some discussion.

Novelty alone in an approach to achieving a goal is not sufficient to warrant constructing a new schema. Consider, for example, a plot from the "Mission Impossible" television series. These plots are noteworthy in that they are very novel. They all use bizarre methods to achieve rather peculiar goals. Furthermore, they are always successful. However, the goals achieved are not the type that arise in ordinary life and the resources and skills needed are so specialized and uncommon that the same solution would never be applicable again.

Clearly a new schema should be constructed only if there is a reasonable expectation that it will be helpful in future processing. If a schema will never be used again, it should not be constructed in the first place.

How can the utility of a particular goal be judged? The answer to this is closely tied to where goals come from. Achieving a goal which arises from general

conditions important to an individual's well-being and using readily available resources is likely to result in an interesting new schema, one which will arise again and again. For the solution, we use an aspect of Schank & Ableson's theory of planning [16]. In their view themes give rise to the highest level goals (goals which are not simply subgoals in the achievement of other goals). Interpersonal and Life themes are what we are interested in. An example of the former is a husband offering (and therefore, at some level, wanting) to type a term paper for his overworked student wife. Examples of the latter are attempting to satisfy one's hunger, to gain money, or to relieve boredom. Life themes give rise to goals that require no further justification. Our example, which demonstrates a new way to gain money, relates directly to a life theme and therefore satisfies this criterion.

Criterion 4 is self-explanatory. The idea is that the system should not bother constructing schemata that are much less efficient than similar already-known schemata. In a natural language input this would occur only if a character were using a highly sub-optimal plan. This criterion is somewhat application dependent. For certain applications one might wish to construct all possible new schemata, even those stemming from bad examples.

The fifth criterion has been discussed elsewhere [2]. As this is a short paper describing on-going research it is not appropriate to repeat it here. Suffice it to say that there is a taxonomy of explanatory acquisition techniques: schema composition, secondary effect elevation, volitionalization, and schema alteration. The technique that is matched has implications for exactly how the generalization is performed.

2.3. The Generalization Process

Assuming the input is completely understood (with data dependency links to inference rules justifying the understanding) and the five tests for learning have all been met, we must now perform the actual generalization. The generalization process consists in replacing the objects and actions in the understood representation with abstract counterparts. These counterparts are the most abstract possible while still preserving the validity of the inference justification network.

Consider again the example of John blackmailing Fred. One proposition that is a part of the understood representation is that Fred decided to pay John \$15,000. This decision event is not mentioned in the story; it is one of the events that must be inferred during understanding to build a causally complete representation. This action is justified to the system by a number of other propositions. Among these supporting propositions are some supplied by the schema DECIDE (which we assume the system already possesses). These inferences from DECIDE are:

- 1) The decider must be at least a higher animate.
- 2) The decider must be capable of a number of alternative possible actions.
- 3) The decider must know what the alternatives are.
- 4) The chosen alternative will be among the most beneficial/least detrimental to the decider.

Thus, these (and other) justifications are tied to the representation of Fred's decision through data dependency links. Fred's decision is believable to the system because Fred, in fact, is a higher animate, he knows at least two alternatives - paying John or losing the \$100,000 and being arrested, and 3) he probably sees losing \$15,000 as less detrimental than losing \$100,000 and going to jail. These

justifications are supplied in the form of pointers to the above inference rules during the understanding procedure.

Now consider the process of generalizing the amount of \$15,000. This money appears several places in the representation. Among them is Fred's decision to give the money to John. The generalization process asks "What are the constraints that the \$15,000 must satisfy?" They are: 1) Fred must possess it or be able to get it and 2) Fred must value it less than the alternative. Thus, the system realizes that values other than \$15,000 would also work. Any amount would work such that 1) Fred can raise that amount and 2) Fred would still prefer giving it up to the alternative. Thus the amount of \$15,000 is replaced by a schema variable with the appropriate binding constraints. Of course, Fred is also generalized. He is replaced by a schema variable that must be a human who has performed an illegal act that is known to a person who is the schema variable replacing John. Thus the binding constraints on the variable that replaces \$15,000 are specified in terms of the variable that replaces Fred. Events are generalized as well. The event of John sending Fred an interoffice memo is generalized to any communication method. The system realizes via the inference justification network that the important effect of that event is to achieve a certain mental state in Fred. The precise method is of no importance. Some elements of the input story are totally discarded. For example, the fact that Fred is John's boss is not mentioned in the inference justification network. It is therefore eliminated the new schema.

Through these and other generalizations the system can construct a first version of a BLACKMAIL schema. The schema might not be perfect. There may be later stories that do not quite fit and require further modification of the schema. However, it is a reasonably general schema that is likely to help a good deal in processing future similar stories.

3. Conclusion

There are several concluding points

1) Unlike many learning system (e.g., [9], [12], [21], [23]) explanatory schema acquisition does not depend on correlational evidence. It is capable of one trial learning, but the learning is not based on analogical reasoning like [20] and [22]. It is somewhat similar to Soloway's view of learning [18]. There is also some resemblance to the MACROPS notion in the STRIPS system [5].

2) The approach is heavily knowledge-based. A great deal of background knowledge must be present for learning to take place. In this respect explanatory schema acquisition follows the current trend in AI learning and discovery systems perhaps traceable to Lenat [11].

3) The learning mechanism is not "failure-driven" as is the MOPs approach ([15], [8], [10]). In that view learning takes place in response to incorrect predictions by the system. In explanatory acquisition learning is usually stimulated by positive inputs which encounter no particular problems or prediction failures.

4) The absolute representation power of the system is not enhanced by learning new schemata. This statement is only superficially surprising. Indeed, Fodor [7] shows that this must be true of all self-consistent learning systems. Explanatory schema acquisition does, however, increase processing efficiency. Since all real-world systems are resource limited, this learning technique does, in fact, increase the system's processing power.

Acknowledgement

I am indebted to the members of the CSL learning group: Chris Debrunner, Paul Harrington, Paul O'Rourke, and Alberto Segre.

REFERENCES

- [1] Cullingford, R. (1978) Script application: computer understanding of newspaper stories. Research Report 116, Yale Computer Science Department, New Haven, CT.
- [2] DeJong, G. (1982) Automatic schema acquisition in a natural language environment. Proceedings of the National Conference on Artificial Intelligence pp. 410-413. Pittsburgh, PA.
- [3] DeJong, G. (1979) Prediction and substantiation: a new approach to natural language processing. Cognitive Science 3, 251-273.
- [4] Doyle, J. (1979) A truth maintenance system. Artificial Intelligence 12(2), pp. 231-272. Technical report TR-419, MIT, Cambridge, MA
- [5] Fikes, R., Hart, P., and Nilsson, N. (1972) Learning and executing generalized robot plans. Artificial Intelligence, 3(2), 251-288.
- [6] Fikes, R. (1975) Deductive retrieval mechanisms for state description models. Proceedings of the Fourth International Joint Conference on Artificial Intelligence. pp. 99-106.
- [7] Fodor, J. (1975) The Language of Thought Thomas Y. Crowell Company, New York
- [8] Kolodner, J. (1980) Retrieval and organizational strategies in conceptual memory: a computer model. Computer science research report 187, Ph. D. dissertation, Yale University, New Haven, CT.
- [9] Langley, P. (1981) Data-driven discovery of physical laws. Cognitive Science 5, 31-54.
- [10] Lebowitz, M., (1980) Generalization and memory in an integrated understanding system. Computer science research report 186, Ph. D. dissertation, Yale University, New Haven, CT.
- [11] Lenat, D., (1976) AM: an artificial intelligence approach to discovery in mathematics as heuristic search. AIM-286, AI Laboratory, Stanford University, Stanford, CA.
- [12] Michalski, R. (1977) A system of programs for computer aided induction: a summary. Proceedings of the Fifth International Joint Conference on Artificial Intelligence pp. 319-320, Cambridge, MA.

- [13] Mitchell, T. (1977) Version spaces: a candidate elimination approach to rule learning. Proceedings of the Fifth International Joint Conference on Artificial Intelligence, pp. 305-310. Cambridge, MA.
- [14] Mitchell, T. (1982) Generalization as search. Artificial Intelligence 18(2), pp. 203-226.
- [15] Schank, R., (1980) Language and memory. Cognitive Science 4, 243-283.
- [16] Schank, R. Abelson, R., (1977) Scripts Plans Goals and Understanding. Lawrence Erlbaum Associates, Hillsdale, NJ.
- [17] Schmidt, C. and Sridharan, N. (1977) Plan recognition using a hypothesize and revise paradigm: an example. Proceedings of the Fifth International Joint Conference on Artificial Intelligence, pp. 480-486.
- [18] Soloway, E., (1977) Knowledge directed learning using multiple levels of description. Ph. D. dissertation, Computer Science Department, University of Massachusetts, Amherst, MA.
- [19] Wilenski, R. (1978) Understanding goal-base stories, Ph.D. dissertation, Yale Computer Science Report 140, September, Yale University, New Haven, CT.
- [20] Winston, P. (1980) Learning and reasoning by analogy, CACM 23(12).
- [21] Winston, P. (1982) Learning new principles from precedents and exercises. Artificial Intelligence 19(3), pp. 321-350.
- [22] Carbonell, J. (1982) Experiential Learning in analogical problem solving. Proceedings of the National Conference on Artificial Intelligence. Pittsburgh, PA.
- [23] Fox, M. and Reddy, R., (1977) Knowledge-guided learning of structural descriptions. Proceedings of the fifth IJCAI, Cambridge, MA.

END

FILMED

10-83

DTIC